



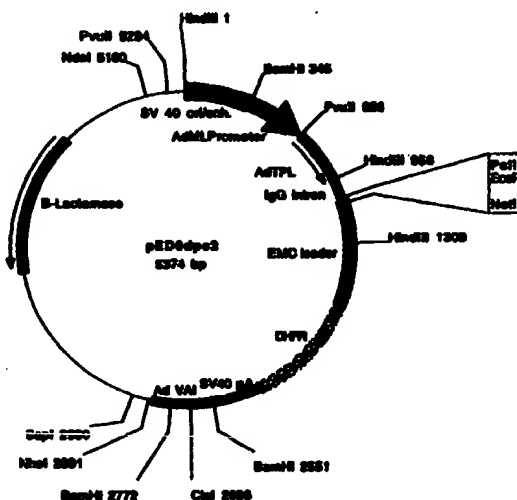
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, A61K 38/17		A2	(11) International Publication Number: WO 98/31802
			(43) International Publication Date: 23 July 1998 (23.07.98)
(21) International Application Number: PCT/US98/01007		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 21 January 1998 (21.01.98)			
(30) Priority Data: 60/072,103 21 January 1997 (21.01.97) US 08/877,035 16 June 1997 (16.06.97) US 09/010,047 20 January 1998 (20.01.98) US			
(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).			
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).		Published Without international search report and to be republished upon receipt of that report.	
(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).			

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED04pc2
Plasmid size: 3374 bp

Comments/References: pED04pc2 is derived from pED04pc1 by insertion of a new polylinker to facilitate cDNA cloning. SET cDNAs are cloned between BamHI and NheI. pED vectors are described in Kaufman et al. (1991), NAR 19: 4488-4492.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NR	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

5

10

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

15

This application is a continuation-in-part of the following applications: Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/786,161), filed January 21, 1997; and Ser. No. 08/877,035, filed June 16, 1997, which is a continuation-in-part of Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/786,161), filed January 21, 1997; all of which are incorporated by reference herein.

20

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

25

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the

30

35

40

cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

5 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID
10 NO:1 from nucleotide 101 to nucleotide 428;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 325 to nucleotide 461;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AA365_1 deposited under accession
15 number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AA365_1 deposited under accession number
20 ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- (i) a polynucleotide encoding a protein comprising a fragment of the
25 amino acid sequence of SEQ ID NO:2 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein
30 of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 101 to nucleotide 428; the nucleotide sequence of SEQ ID NO:1

from nucleotide 325 to nucleotide 461; the nucleotide sequence of the full-length protein coding sequence of clone AA365_1 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone AA365_1 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 76 to amino acid 109.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
- (b) the amino acid sequence of SEQ ID NO:2 from amino acid 76 to amino acid 109;
- (c) fragments of the amino acid sequence of SEQ ID NO:2; and
- (d) the amino acid sequence encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 76 to amino acid 109.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 245 to nucleotide 421;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 413 to nucleotide 421;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BL67_2 deposited under accession number ATCC 98296;

- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BL67_2 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BL67_2 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BL67_2 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 245 to nucleotide 421; the nucleotide sequence of SEQ ID NO:3 from nucleotide 413 to nucleotide 421; the nucleotide sequence of the full-length protein coding sequence of clone BL67_2 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone BL67_2 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BL67_2 deposited under accession number ATCC 98296. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 20.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3 or SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 20;

(c) fragments of the amino acid sequence of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone

5 BL67_2 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 20.

10 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 292 to nucleotide 468;

15 (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BN130_1 deposited under accession number ATCC 98296;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BN130_1 deposited under accession number ATCC 98296;

20 (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BN130_1 deposited under accession number ATCC 98296;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BN130_1 deposited under accession number ATCC 98296;

25 (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

30 (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:6 from nucleotide 292 to nucleotide 468; the nucleotide sequence of the full-length protein coding sequence of clone BN130_1 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone BN130_1 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BN130_1 deposited under accession number ATCC 98296.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:6 or SEQ ID NO:8.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:7;
- (b) fragments of the amino acid sequence of SEQ ID NO:7; and
- (c) the amino acid sequence encoded by the cDNA insert of clone BN130_1 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:7.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 137 to nucleotide 2146;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 218 to nucleotide 2146;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 189 to nucleotide 397;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CG99_2 deposited under accession number ATCC 98296;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CG99_2 deposited under accession number ATCC 98296;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CG99_2 deposited under accession number ATCC 98296;

5 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CG99_2 deposited under accession number ATCC 98296;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity;

10 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

15 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 137 to nucleotide 2146; the nucleotide sequence of SEQ ID NO:9 from nucleotide 218 to nucleotide 2146; the nucleotide sequence of SEQ ID NO:9 from nucleotide 189 to nucleotide 397; the nucleotide sequence of the full-length protein coding
20 sequence of clone CG99_2 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone CG99_2 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CG99_2 deposited under accession number ATCC 98296. In yet other preferred
25 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 50 to amino acid 87.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

30 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 50 to amino acid 87;

(c) fragments of the amino acid sequence of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone
5 CG99_2 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 50 to amino acid 87.

In one embodiment, the present invention provides a composition comprising an
10 isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 123 to nucleotide 1247;

15 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 1 to nucleotide 419;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO618_1 deposited under accession number ATCC 98296;

20 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO618_1 deposited under accession number ATCC 98296;

25 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the
30 amino acid sequence of SEQ ID NO:12 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 123 to nucleotide 1247; the nucleotide sequence of SEQ ID NO:11 from nucleotide 1 to nucleotide 419; the nucleotide sequence of the full-length protein coding sequence of clone CO618_1 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone CO618_1 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 99.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:12;
 - (b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 99;
 - (c) fragments of the amino acid sequence of SEQ ID NO:12; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 99.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 208 to nucleotide 402;

- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 184 to nucleotide 300;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO629_2 deposited under accession number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO629_2 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 208 to nucleotide 402; the nucleotide sequence of SEQ ID NO:13 from nucleotide 184 to nucleotide 300; the nucleotide sequence of the full-length protein coding sequence of clone CO629_2 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone CO629_2 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 31.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group

5 consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 31;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- 10 (d) the amino acid sequence encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 31.

15 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID
- 20 NO:15 from nucleotide 94 to nucleotide 1059;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 387;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT645_1 deposited under accession
- 25 number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT645_1 deposited under accession number
- 30 ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

5 (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 94 to nucleotide 1059; the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 387; the nucleotide sequence of the full-length protein coding sequence of clone CT645_1 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone CT645_1 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 98.

20 Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

25 (a) the amino acid sequence of SEQ ID NO:16;
(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 98;

(c) fragments of the amino acid sequence of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296;

30 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 98.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;
- 5 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 640 to nucleotide 1029;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 611 to nucleotide 814;
- 10 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CW383_1 deposited under accession number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature
15 protein coding sequence of clone CW383_1 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid
20 sequence of SEQ ID NO:18;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- 25 (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID
30 NO:17 from nucleotide 640 to nucleotide 1029; the nucleotide sequence of SEQ ID NO:17 from nucleotide 611 to nucleotide 814; the nucleotide sequence of the full-length protein coding sequence of clone CW383_1 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone CW383_1 deposited under accession number ATCC 98296. In other preferred embodiments, the

polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid

5 63.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group

10 consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 63;
- (c) fragments of the amino acid sequence of SEQ ID NO:18; and
- 15 (d) the amino acid sequence encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 63.

20 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID
- 25 NO:19 from nucleotide 509 to nucleotide 718;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN167_1 deposited under accession number ATCC 98296;
- (d) a polynucleotide encoding the full-length protein encoded by the
- 30 cDNA insert of clone DN167_1 deposited under accession number ATCC 98296;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN167_1 deposited under accession number ATCC 98296;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN167_1 deposited under accession number ATCC 98296;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

5 (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

10 (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 509 to nucleotide 718; the nucleotide sequence of the full-length
15 protein coding sequence of clone DN167_1 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone DN167_1 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DN167_1 deposited under accession number ATCC 98296. In
20 yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 65.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

25 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;
(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to
30 amino acid 65;

(c) fragments of the amino acid sequence of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone DN167_1 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 65.

In one embodiment, the present invention provides a composition comprising an
5 isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 177 to nucleotide 470;
- 10 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 237 to nucleotide 470;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN711_2 deposited under accession number ATCC 98296;
- 15 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN711_2 deposited under accession number ATCC 98296;
- 20 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22;
- (i) a polynucleotide encoding a protein comprising a fragment of the
25 amino acid sequence of SEQ ID NO:22 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- 30 (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:21 from nucleotide 177 to nucleotide 470; the nucleotide sequence of SEQ ID NO:21 from nucleotide 237 to nucleotide 470; the nucleotide sequence of the full-length protein

coding sequence of clone DN711_2 deposited under accession number ATCC 98296; or the nucleotide sequence of the mature protein coding sequence of clone DN711_2 deposited under accession number ATCC 98296. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296. In yet other preferred
5 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22 from amino acid 1 to amino acid 94.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ
10 ID NO:21.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:22;
- 15 (b) the amino acid sequence of SEQ ID NO:22 from amino acid 1 to amino acid 94;
- (c) fragments of the amino acid sequence of SEQ ID NO:22; and
- (d) the amino acid sequence encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296;
- 20 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:22 or the amino acid sequence of SEQ ID NO:22 from amino acid 1 to amino acid 94.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial,
25 yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

- 30 (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and
- (b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a
5 pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a
10 pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

15

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide
20 sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its
25 sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal
30 sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "AA365_1"

A polynucleotide of the present invention has been identified as clone "AA365_1". AA365_1 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was
5 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AA365_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AA365_1 protein").

The nucleotide sequence of AA365_1 as presently determined is reported in SEQ
10 ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the AA365_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Another possible reading frame and predicted amino acid sequence encoded by AA365_1 is encoded by base pairs 157-610 of SEQ ID NO:1 and is reported in SEQ ID NO:33. Amino acids 9 to 21 of SEQ ID NO:33 are
15 a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 22 of SEQ ID NO:33, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AA365_1 should be approximately 700 bp.

The nucleotide sequence disclosed herein for AA365_1 was searched against the
20 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AA365_1 demonstrated at least some similarity with sequences identified as AA019239 (ze56c09.s1 Soares retina N2b4HR Homo sapiens cDNA clone 362992 3'). Based upon sequence similarity, AA365_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts
25 two potential transmembrane domains within the AA365_1 protein sequence centered around amino acids 33 and 73 of SEQ ID NO:2, respectively; amino acids 14 to 26 are also a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 27 of SEQ ID NO:2.

30 Clone "BL67_2"

A polynucleotide of the present invention has been identified as clone "BL67_2". BL67_2 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. BL67_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BL67_2 protein").

5 The nucleotide sequence of the 5' portion of BL67_2 as presently determined is reported in SEQ ID NO:3. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:4. The predicted amino acid sequence of the BL67_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 44 to 56 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 57, or are a
10 transmembrane domain. Additional nucleotide sequence from the 3' portion of BL67_2, including the polyA tail, is reported in SEQ ID NO:5.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BL67_2 should be approximately 970 bp.

15 The nucleotide sequence disclosed herein for BL67_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. No hits were found in the database.

Clone "BN130_1"

A polynucleotide of the present invention has been identified as clone "BN130_1".
20 BN130_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BN130_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein
25 as "BN130_1 protein").

The nucleotide sequence of the 5' portion of BN130_1 as presently determined is reported in SEQ ID NO:6. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:7. The predicted amino acid sequence of the BN130_1 protein corresponding to the foregoing nucleotide sequence is reported in
30 SEQ ID NO:7. Additional nucleotide sequence from the 3' portion of BN130_1, including the polyA tail, is reported in SEQ ID NO:8.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BN130_1 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for BN130_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BN130_1 demonstrated at least some similarity with sequences identified as AA023887 (mh93g07.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA clone 458556 5'), T99892 (ye68f02.r1 Homo sapiens cDNA clone 122907 5'), U65410 (Human Mad2 (hsMAD2) mRNA, complete cds), and Z13859 (H. sapiens partial cDNA sequence; clone 81C04; strand (-), single read). The predicted amino acid sequence disclosed herein for BN130_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BN130_1 protein demonstrated at least some similarity to sequences identified as D37934 (Rat mRNA for 5E5 antigen, complete cds). Based upon sequence similarity, BN130_1 proteins and each similar protein or peptide may share at least some activity.

Clone "CG99_2"

A polynucleotide of the present invention has been identified as clone "CG99_2". CG99_2 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CG99_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CG99_2 protein").

The nucleotide sequence of CG99_2 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CG99_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Amino acids 15 to 27 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 28, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CG99_2 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for CG99_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CG99_2 demonstrated at least some similarity with sequences identified as L03813 Rattus norvegicus neurotrophin-3 receptor (trkC) mRNA, complete

cds). Based upon sequence similarity, CG99_2 proteins and each similar protein or peptide may share at least some activity.

Clone "CO618_1"

5 A polynucleotide of the present invention has been identified as clone "CO618_1". CO618_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO618_1 is a full-length
10 clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO618_1 protein").

The nucleotide sequence of CO618_1 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CO618_1 protein corresponding to the foregoing
15 nucleotide sequence is reported in SEQ ID NO:12.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO618_1 should be approximately 1750 bp.

The nucleotide sequence disclosed herein for CO618_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and
20 FASTA search protocols. CO618_1 demonstrated at least some similarity with sequences identified as D87074 (Human male bone marrow myeloblast mRNA for KIAA0237 protein, complete cds), H34957 (EST110575 Rattus sp. cDNA), and T33172 (EST56933 Homo sapiens cDNA 3' end similar to None). The predicted amino acid sequence disclosed herein for CO618_1 was searched against the GenPept and GeneSeq amino acid
25 sequence databases using the BLASTX search protocol. The predicted CO618_1 protein demonstrated at least some similarity to sequences identified as AF007836 (rab3 effector [Rattus norvegicus]), D87074 (KIAA0237 protein; similar to a C.elegans protein encoded in cosmid T10A3(U41035) [Homo sapiens]), and D90195 (polyprotein [Japanese encephalitis virus]). CO618_1 protein showed at least some similarity to rat Rim, a
30 putative Rab3 effector. Rim is composed of an amino-terminal zinc-finger motif and carboxy-terminal PDZ and C2 domains. It binds only to GTP (but not to GDP)-complexed Rab3, and interacts with no other Rab protein. Rim is localized to presynaptic active zones in conventional synapses, and to presynaptic ribbons in ribbon synapses (Wang *et al.*, 1997, *Nature* 388(6642): 593-598, incorporated by reference herein). Based upon

sequence similarity, CO618_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the CO618_1 protein sequence centered around amino acid 350 of SEQ ID NO:12.

5

Clone "CO629_2"

A polynucleotide of the present invention has been identified as clone "CO629_2". CO629_2 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was
10 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO629_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO629_2 protein").

The nucleotide sequence of CO629_2 as presently determined is reported in SEQ
15 ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CO629_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Another possible reading frame and predicted amino acid sequence encoded by CO629_2 is encoded by basepairs 417 to 632 of SEQ ID NO:13 and is reported in SEQ ID NO:34; this alternative open reading frame
20 could be joined to the reading frame reported in SEQ ID NO:14 if an insertion or deletion resulting in a frameshift was made in the sequence of SEQ ID NO:13. The TopPredII computer program predicts a potential transmembrane domains within the alternative CO629_2 protein sequence centered around amino acid 50 of SEQ ID NO:34.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone
25 CO629_2 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for CO629_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO629_2 demonstrated at least some similarity with sequences identified as U78857 (*Rattus norvegicus* calcium calmodulin dependent kinase CPG16
30 (cpg16) mRNA, complete cds) and W77588 (me67b11.r1 Soares mouse embryo NbME13.5 14.5 *Mus musculus* cDNA). The predicted amino acid sequence disclosed herein for CO629_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO629_2 protein demonstrated at least some similarity to sequences identified as U78857 (calcium calmodulin dependent kinase

CPG16 [Rattus norvegicus]). Based upon sequence similarity, CO629_2 proteins and each similar protein or peptide may share at least some activity.

Clone "CT645_1"

5 A polynucleotide of the present invention has been identified as clone "CT645_1". CT645_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CT645_1 is a full-length
10 clone, including the entire coding sequence of a secreted protein (also referred to herein as "CT645_1 protein").

The nucleotide sequence of CT645_1 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CT645_1 protein corresponding to the foregoing
15 nucleotide sequence is reported in SEQ ID NO:16.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CT645_1 should be approximately 2560 bp.

The nucleotide sequence disclosed herein for CT645_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and
20 FASTA search protocols. CT645_1 demonstrated at least some similarity with sequences identified as AA430737 (zw20g12.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 769894 5'), H34543 (EST111568 Rattus sp. cDNA 5' end), and N34881 (yy46g12.s1 Homo sapiens cDNA clone 276646 3' similar to contains element MER37 repetitive element). Based upon sequence similarity, CT645_1 proteins and each similar protein or
25 peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the CT645_1 protein sequence centered around amino acids 75 and 275 of SEQ ID NO:16, respectively; amino acids 92 to 104 of SEQ ID NO:16 are a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 105. The nucleotide sequence of CT645_1 indicates that
30 it may contain an A1u repetitive element.

Clone "CW383_1"

A polynucleotide of the present invention has been identified as clone "CW383_1". CW383_1 was isolated from a human fetal brain cDNA library using methods which are

selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CW383_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein
5 as "CW383_1 protein").

The nucleotide sequence of CW383_1 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CW383_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18.

10 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CW383_1 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for CW383_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CW383_1 demonstrated at least some similarity with sequences
15 identified as AA037470 (zk36g04.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 484950 3'), AA037478 (zk36h05.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 484953 3'), AA088729 (zl89c09.r1 Stratagene colon (#937204) Homo sapiens cDNA clone 511792 5'), AA098551 (mk18f12.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 493295 5'), R75968 (yi62d07.r1 Homo sapiens cDNA clone 143821 5'), and
20 X85681 (H.sapiens mRNA for expressed sequence tag, clone CAM tEST1F6 (B)). Based upon sequence similarity, CW383_1 proteins and each similar protein or peptide may share at least some activity.

Clone "DN167_1"

25 A polynucleotide of the present invention has been identified as clone "DN167_1". DN167_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DN167_1 is a full-length
30 clone, including the entire coding sequence of a secreted protein (also referred to herein as "DN167_1 protein").

The nucleotide sequence of DN167_1 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the

predicted amino acid sequence of the DN167_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DN167_1 should be approximately 850 bp.

5 The nucleotide sequence disclosed herein for DN167_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DN167_1 demonstrated at least some similarity with sequences identified as AA405121 (zu65d01.r1 Soares testis NHT Homo sapiens cDNA clone 742849 5'), AC000045 (00064; HTGS phase 3, complete sequence), R51018 (yg71b02.s1 Homo sapiens cDNA clone 38639 3' similar to contains Alu repetitive element; contains MSR1 repetitive element), and Z70233 (Human DNA sequence from cosmid V870H8, between markers DXS366 and DXS87 on chromosome X contains ESTs). Based upon sequence similarity, DN167_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of DN167_1 indicates that it may contain an Alu
10 repetitive element.
15

Clone "DN711_2"

A polynucleotide of the present invention has been identified as clone "DN711_2". DN711_2 was isolated from a human fetal brain cDNA library using methods which are
20 selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DN711_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DN711_2 protein").

25 The nucleotide sequence of DN711_2 as presently determined is reported in SEQ ID NO:21. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DN711_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:22. Amino acids 8 to 20 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at
30 amino acid 21, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DN711_2 should be approximately 700 bp.

The nucleotide sequence disclosed herein for DN711_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. DN711_2 demonstrated at least some similarity with sequences identified as AA127018 (zl22a03.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 502636 3'), AA128221 (zl22a03.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 502636 5'), AA444139 (zv51f01.r1 Soares testis NHT Homo sapiens cDNA clone 757177 5'), AA460975 (zx63c11.s1 Soares total fetus Nb2HF8 9w Homo sapiens), D62671 (Human aorta cDNA 5'-end GEN-312F07), D79837 (Human aorta cDNA 5'-end GEN-335E12), D79860 (Human aorta cDNA 5'-end GEN-342C02), R24595 (yg36h06.r1 Homo sapiens cDNA clone 34631 5'), and R40285 (yf80h03.s1 Homo sapiens cDNA clone). Based upon sequence similarity, DN711_2 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones AA365_1, BL67_2, BN130_1, CG99_2, CO618_1, CO629_2, CT645_1, CW383_1, DN167_1, and DN711_2 were deposited on January 17, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98296, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

	<u>Clone</u>	<u>Probe Sequence</u>
10	AA365_1	SEQ ID NO:23
	BL67_2	SEQ ID NO:24
	BN130_1	SEQ ID NO:25
	CG99_2	SEQ ID NO:26
	CO618_1	SEQ ID NO:27
15	CO629_2	SEQ ID NO:28
	CT645_1	SEQ ID NO:29
	CW383_1	SEQ ID NO:30
	DN167_1	SEQ ID NO:31
	DN711_2	SEQ ID NO:32

20

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

25

The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
- 30 (b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with γ -³²P ATP (specific activity 6000 Ci/mmmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated

label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4×10^6 dpm/pmole.

- 5 The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 μ l of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 μ g/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the
- 10 dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 μ g/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

- Standard colony hybridization procedures should then be used to transfer the
- 15 colonies to nitrocellulose filters and lyse, denature and bake them.

- The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 μ g/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at
- 20 a concentration greater than or equal to 1×10^6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The
- 25 filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

- The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis,
- 30 hybridization analysis, or DNA sequencing.

 Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S.

McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* 15(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* 62(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* 58: 1-

39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided.

5 Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated,

10 through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* 14(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91(2): 719-722;

15 all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably

20 are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor),

25 the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

30 Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing

the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that
5 shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or
10 polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides .

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

20 The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as
25 stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) [‡]	Hybridization Temperature and Buffer [†]	Wash Temperature and Buffer [†]
5	A	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
	B	<50	T _B [*] ; 1xSSC	T _B [*] ; 1xSSC
	C	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
	D	<50	T _D [*] ; 1xSSC	T _D [*] ; 1xSSC
	E	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
	F	<50	T _F [*] ; 1xSSC	T _F [*] ; 1xSSC
10	G	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
	H	<50	T _H [*] ; 4xSSC	T _H [*] ; 4xSSC
	I	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
	J	<50	T _J [*] ; 4xSSC	T _J [*] ; 4xSSC
	K	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
	L	<50	T _L [*] ; 2xSSC	T _L [*] ; 2xSSC
15	M	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
	N	<50	T _N [*] ; 6xSSC	T _N [*] ; 6xSSC
	O	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
	P	<50	T _P [*] ; 6xSSC	T _P [*] ; 6xSSC
	Q	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
	R	<50	T _R [*] ; 4xSSC	T _R [*] ; 4xSSC

[‡] The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

[†] SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

^{*}T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6(log₁₀[Na⁺]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na⁺] is the concentration of sodium ions in the hybridization buffer ([Na⁺] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds.,
5 John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or
10 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an
15 expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably
20 linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the
25 protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

30 Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methy1 or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance
5 with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

10 The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith,
15 including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally
20 provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another
25 amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be
30 expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured
5 by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter
10 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node
15 cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic
20 cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature*
25 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols*
30 *in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient
5 by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic
10 acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function
15 (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides.
20 For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used
25 to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II
30 molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

10 The activity of a protein of the invention may, among other means, be measured by the following methods:

 Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, *Immunologic studies in Humans*); Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bowman et al., *J. Virology* 61:1992-1998; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnoli et al., *Cellular Immunology* 133:327-341, 1991; Brown et al., *J. Immunol.* 153:3079-3092, 1994.

 Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J. Immunol.* 144:3028-3033, 1990; and *Assays for B cell function: In vitro antibody production*, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

30 Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnoli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of

5 hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or

10 *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

15 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood

20 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359,

25 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

30

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

5 A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

10 A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of
15 congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

 A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce
20 differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

25 Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and
30 other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce
5 differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in
10 the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve
15 tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present
20 invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of
25 non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac)
30 and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting
5 differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described
10 in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in:
Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year
15 Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related
20 activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals
25 and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example,
30 United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., *Endocrinology* 91:562-572, 1972; Ling et al., *Nature* 321:779-782, 1986; Vale et al., *Nature* 321:776-779, 1986; Mason et al., *Nature* 318:659-663, 1985; Forage et al., *Proc. Natl. Acad. Sci. USA* 83:3091-3095, 1986.

5

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells.

- 10 Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses
- 15 against the tumor or infecting agent.

- A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population
- 20 of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

- Assays for chemotactic activity (which will identify proteins that induce or prevent
- 25 chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene
- 30 Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. *J. Clin. Invest.* 95:1370-1376, 1995; Lind et al. *APMIS* 103:140-146, 1995; Muller et al. *Eur. J. Immunol.* 25: 1744-1748; Gruber et al. *J. of Immunol.* 152:5860-5867, 1994; Johnston et al. *J. of Immunol.* 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

10 The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 15 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

30 The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 5 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in 10 the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat 15 inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting 20 from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major 25 roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

30 The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

5 E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to
10 their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention
15 encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue
20 in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and
25 polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the
30 cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides
5 encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

10 Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or
15 tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

20 Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height,
25 weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein,
30 carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen

5 in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

10 A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term

15 "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11,

20 IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention,

25 or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

30 A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The

5 pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone.

10 Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not
15 increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μ g to about 100 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1 μ g to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the
20 present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous
25 therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the
30 carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting
5 and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When
10 administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also
15 optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the
20 developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular
25 application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins
30 or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns.

- 5 In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose,
10 ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20
15 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

- 20 In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), and insulin-like growth factor (IGF).

- 25 The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering
30 various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline
5 labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without
10 limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

15 Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
- (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
- (iii) NUMBER OF SEQUENCES: 34
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Genetics Institute, Inc.
 - (B) STREET: 87 CambridgePark Drive
 - (C) CITY: Cambridge
 - (D) STATE: MA
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Sprunger, Suzanne A.
 - (B) REGISTRATION NUMBER: 41,323
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (617) 498-8284
 - (B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 633 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

CTCTTACCCC TCCTCCTGTT TCATGGACAT TTCAGTGTG CCAGTGAAAA AGTAATGGTT      60
TTAATTTGGT CCTTATTTT AACCTGCCCC TGAGACTTAT ATGCTTGTTT ATACCATGTA      120
CGTAGTGTGT GATTGTATGT GTTGTATTT GTCCACATGT CCCAAACAT GGGCTGTTAC      180
TTCCTTTTTC TATCTTGGTT TCCTTATTC CACCCTTCTT TTTCCACCCA GGTATCTGGA      240
CAGGAAGACT TCTCCCATCA GCTTTACCAG AGGAAGCTGC AGGCCCCACT GTGGCCCAGC      300
TCCCTGGGCA TCACTGATTG CTGTCAGTAT GTCACCTCCT GTCACCCCAA GAGATCAGAG      360
AGACGCAAGT ATGGCCGAGA CTTCTGCTA CGTTTCCGCT TCTGCAGCAT CGCTTGTCAG      420
CGACCAGTAG GACTGGTCCT TATGGAAGGA GTGACAGATA CTAAGCCAGA GCGACCTGCG      480
GGTGGGCTG AGTCTGTCCT TGAGGAAGAT GCATCGGAGC TTGAGCCTGC CTTCTCCAGG      540
ACTGTAGGTA CCATCCAGCA CTGCCCTCAC CTGACGTCAG TATATACCCC CTCTACCCCC      600
TCTAGAGCCA AAAAAAAAAA AAAAAAAAAA AAA                                     633

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Leu Val Tyr Thr Met Tyr Val Val Cys Asp Cys Met Cys Leu Tyr
1           5           10          15

Leu Ser Thr Cys Pro Lys Thr Trp Ala Val Thr Ser Phe Phe Tyr Leu
20          25          30

Gly Phe Leu Ile Pro Thr Leu Leu Phe Pro Pro Arg Tyr Leu Asp Arg
35          40          45

Lys Thr Ser Pro Ile Ser Phe Thr Arg Gly Ser Cys Arg Pro His Cys
50          55          60

```

Gly Pro Ala Pro Trp Ala Ser Leu Ile Ala Val Ser Met Ser Pro Pro
 65 70 75 80

Val Thr Pro Arg Asp Gln Arg Asp Ala Ser Met Ala Glu Thr Ser Cys
 85 90 95

Tyr Val Ser Ala Ser Ala Ala Ser Leu Val Ser Asp Gln
 100 105

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 421 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGGCCGCAG GTCTAGAATT CAATCGGGGA AGTGGGCTGT CCAGCTGATG TGCAGCCTCA 60

AAAGAGACCC CAAGCCAGAC CCACCCACGA AAGGTACCCC TGAATTCTGG ACTCATAAAA 120

TCTCTTAGAT AATACATATT TGTGTGTTTA ACCCACTCAG TTTTGCATTG GTTACACAGC 180

AATAGAAAAT AAAGGAGGCA GATGAAATTG ATGGAGAACA ATTGCAACGA AAACAGAATA 240

CACAATGCAC GAGCCTGTGT CAGGAATGAC AGTGCAATCC ACGGAAGAGT TGCACAGAGA 300

GAGACCAGTT ACAAGATGCC ATCAGGATAA CGGGAGCTGC GCTGGGCAGC TGGGCCTTCC 360

GACATCCTCC CTGGTTGCGG TGGCAGTGGT GGTGTCTGCT CCAAGTCAAG GCTGGCAGAG 420

A 421

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	His	Glu	Pro	Val	Ser	Gly	Met	Thr	Val	His	Ser	Thr	Glu	Glu	Leu
1				5					10					15	
His	Arg	Glu	Arg	Pro	Val	Thr	Arg	Cys	His	Gln	Asp	Asn	Gly	Ser	Cys
			20					25					30		
Ala	Gly	Gln	Leu	Gly	Leu	Pro	Thr	Ser	Ser	Leu	Val	Ala	Val	Ala	Val
		35					40					45			
Val	Val	Ser	Ala	Pro	Ser	Gln	Gly	Trp	Gln	Arg					
	50					55									

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 393 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TCTGGAACA	CAAACCTCAC	TGAGTAACAA	AGAAACAGTA	TTTGAGGAAA	GCCAATGTGA	60
TGAAAGAGAG	ACGCCAGACT	GGAAGGAGGG	ACGGTCATCT	GAGAACATAC	AAAGCAAGTG	120
GCAAGCGTTA	GTTTACACGT	CCAAGGTGAC	CGTCACATGC	TCTGTCCAGG	CAAGAGAGGC	180
AAGTGGACAC	CCTGAGAAGG	AAAAATGTTG	TCAAATATAA	AACCTAATAT	GTGAACCTAG	240
GAGTAAATG	GACACAGCTT	GCGAGCAAAG	AGACAGTGAG	AAGATGAAGC	CAAGGCACTC	300
TCCAGAATGC	AGGTGACGGG	AAAAAGAGAC	AGTAACAGAT	CCAAAGAGGT	CCAGAAGGAA	360
TGACCATGGA	GCGGAAAGAA	AGAAAAAAAAA	AAA			393

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 468 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCTCCGGTGG CTGCGCAGCG TCGGCGGGGA GCGCGGGCTG CGGAGAGGCG GGCCGGGGCCA 60
 AGCGGAGCCG AGCGAGCGGG AGCGCGGCGT CCGGGAGGCG GCGGAGACGC GGGGCTCGGA 120
 GGGTCAGCCT CTTATCGTAG CAGGTCTCCT CGGCACGCCC CCCTTGTTTC GCCCCACGGC 180
 CAAGCCCGCC GCGGGCCGGC GTGCGCTGGT CACTGAGGCC CAGGTCGCCG CCGCGGCGCG 240
 TTTTGGAAAT CATGAATCCT GTTTATAGTC CTGGATCTTC TGGGGTCCC TATGCAAATG 300
 CCAAAGGAAT TGGTTATCCA GCTGGTTTTTC CCATGGGCTA TGCAGCAGCA GCTCCCTGCC 360
 TATTCTCCTA ACATGTATCC TGGAGCGAAT CCTACCTTCC AAACAGGTTA CACTCCTGGC 420
 ACACCTTACA AAGTGTCTTG TTCCCCACC AGCGGGGCTG TGCCACCG 468

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Pro	Lys	Glu	Leu	Val	Ile	Gln	Leu	Val	Phe	Pro	Trp	Ala	Met	Gln
1				5				10					15		
Gln	Gln	Leu	Pro	Ala	Tyr	Ser	Pro	Asn	Met	Tyr	Pro	Gly	Ala	Asn	Pro
			20					25					30		
Thr	Phe	Gln	Thr	Gly	Tyr	Thr	Pro	Gly	Thr	Pro	Tyr	Lys	Val	Ser	Cys
		35					40					45			
Ser	Pro	Thr	Ser	Gly	Ala	Val	Pro	Pro							
		50					55								

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 366 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

AAGTACCGGG CCCAGGAAC GCCCACTTAC AGCTATGTGC CCCCTCAGTG GTGATCACCT      60
GCAAATGTTT GAGGACGGAG CTGTGCAGTC ACATTATTGG GGATTCCACA GCTGGTGCTG      120
CAGGCCTTGC GCCTCCAACC AGGACTTTCT TCTTAATGCT CTCGACACTT AGCTAAACAC      180
GACTATATCC CGGCCAGCA GGCCCCAGCG CCGTTAGTCT CCAGCTGACT CTGTGGGTG      240
GTCTTAAAGC AAATCTGTT TTGTGGACTG CCTGGCAATT TTTTAGCTAA CTGTAATGAT      300
AAAAAGGGAG TATTAATCTA TTCTGAATCA TATCTAGTTG AATGCATGTT TAAAAAAAAA      360
AAAAAA                                         366

```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2423 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

AGGGCCCGCC GCTGGGATGC CGAGCGCCCG CGCCGCCGCT GCCTCTGTCC TCCGCGCGCT      60
GCTCAGCTGA AGGCGCACAG GATTCAATTA CTGGACTTGT CAACTCTGCC AGTGACGTG      120
CCATTTCTCT TCCACTATGA GAGGACCGAT TGTATTGCAC ATTTGTCTGG CTTTCTGTAG      180
CCTTCTGCTT TTCAGCGTTG CCACACAATG TCTGGCCTTC CCCAAAATAG AAAGGAGGAG      240
GGAGATAGCA CATGTTTCATG CGGAAAAAGG GCAGTCCGAT AAGATGAACA CCGATGACCT      300
AGAAAATAGC TCTGTTACCT CAAAGCAGAC TCCCCAACTG GTGGTCTCTG AAGATCCAAT      360
GATGATGTCA GCAGTACCAT CGGCAACATC ATTAAATAAA GCATTCTCGA TTAACAAAGA      420
AATCCAGCCT GGACAAGCTG GGCTCATGCA AACAGAACGC CCTGGTGTTC CCACACCTAC      480
TGAGTCAGGT GTCCCTCAG CTGAAGAAGT ATTTGGTTCC AGCCAGCCAG AGAGAATATC      540
TCCTGAAAGT GGACTTGCCA AGGCCATGTT AACCATTGCT ATCACTGCGA CTCCTTCTCT      600
GACTGTTGAT GAAAAGGAGG AACTCCTTAC AAGCACTAAC TTTAGCCCA TTGTAGAAGA      660
GATCACAGAA ACCACAAAAG GTTTTCTGAA GTATATGGAT AATCAATCAT TTGCAACTGA      720

```

AAGTCAGGAA GGAGTTGGTT TGGGACATTC ACCTTCATCC TATGTGAATA CTAAGGAAAT	780
GCTAACCACC AATCCAAAGA CTGAGAAATT TGAAGCAGAC ACAGACCACA GGACAACTTC	840
TTTTCTGGT GCTGAGTCCA CAGCAGGCAG TGAGCCTGGA AGCCTCACCC CTGATAAGGA	900
GAAGCCTTCG CAGATGACAG CTGATAACAC CCAGGCTGCT GCCACCAAGC AACCCTCGA	960
AACTTCCGAG TACACCCTGA GTGTTGAGCC AGAACTGAT AGTCTGCTGG GAGCCCCAGA	1020
AGTCACAGTG AGTGTGAGCA CAGCTGTTCC AGCTGCCTCT GCCTTAAGTG ATGAGTGGGA	1080
TGACACCAAA TTAGAGAGTG TAAGCCGGAT AAGGACCCCC AAGCTTGGAG ACAATGAAGA	1140
GACTCAGGTG AGAACGGAGA TGTCTCAGAC AGCACAAGTA AGCCATGAGG GTATGGAAGG	1200
AGGCCAGCCT TGGACAGAGG CTGCACAGGT GGCTCTGGGG CTGCCTGAAG GGGAAACACA	1260
CACGGGCACA GCCCTGCTAA TAGCGCATGG GAATGAGAGA TCACCTGCTT TCACTGATCA	1320
AAGTTCTTTT ACCCCCACAA GTCTGATGGA AGACATGAAA GTTTCCATTG TGAACCTTGCT	1380
CCAAAGTACG GGAGACTTCA CGGAATCCAC CAAGGAAAAC GATGCCCTGT TTTTCTTAGA	1440
AACCCTGTT TCTGTCTCTG TATATGAGTC TGAGGCAGAC CAACTGTTGG GAAATACAAT	1500
GAAAGACATC ATCACTCAAG AGATGACAAC AGCTGTTCAA GAGCCAGATG CCACTTTATC	1560
CATGGTGACA CAAGAGCAGG TTGCTACCCT CGAGCTTATC AGAGACAGTG GCAAGACTGA	1620
GGAAGAAAAG GAGGACCCCT CTCCTGTGTC TGACGTTCTT GGTGTTACTC AGCTGTCAAG	1680
AAGATGGGAG CCTCTGGCCA CTACAATTTT AACTACAGTC GTCCCTTTGT CTTTGAAGT	1740
TACTCCCACT GTGGAAGAAC AAATGGACAC AGTCACAGGG CCAAATGAGG AGTTCACACC	1800
AGTTCTGGGA TCTCCAGTGA CACCTCCTGG AATAATGGTG GGGGAACCCA GCATTTCCCC	1860
TGCACTTCCT GCTTGGAGG CATCCTCTGA GAGAAGAACT GTTGTTCAT CTATTACTCG	1920
TGTTAATACA GCTGCCTCAT ATGGCCTGGA CCAACTTGAA TCTGAAGAGA CGGGGTTTCA	1980
CCATGTTGCC CAGGCTCGTC TCAAACCTCT GGGCTTAAGA AGTCTGCCTG CCTCGGCCTC	2040
CCAAAGTGTT GGAATTACAA GTGTGAACTC CTGCACCCAG CCAAGAAAAT ATTTAAACAG	2100
CTGCCTGAAA TGGAAATTGA ATCCAAAACA TTTTGCCACT GGTGTTTGAA GTGGACTATC	2160
TTGTCTCAAG CCTTTTATTT TTCTCATGAA TGTACCTTCT AAATATTGTT AAGGCTACAA	2220
TGGAGGAAAT TTAAATTAT TATATTTACT ATTTTACAGT GAGTGGAAAA TTTGAAATTT	2280
AGTAAAAAGT TTCAAGAAA TTTTATAGA TTAATAATGG AGTTTAAAAA GTCATGGCGA	2340
TGAGGCATTT AACAAGGAGA GAACTGAATA AATTAACAGA CATGTTTAAG GTAAAAAAA	2400

AAAAAAAAAA AAAAAAAAAA AAA

2423

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 670 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Arg Gly Pro Ile Val Leu His Ile Cys Leu Ala Phe Cys Ser Leu
1           5           10           15

Leu Leu Phe Ser Val Ala Thr Gln Cys Leu Ala Phe Pro Lys Ile Glu
          20           25           30

Arg Arg Arg Glu Ile Ala His Val His Ala Glu Lys Gly Gln Ser Asp
          35           40           45

Lys Met Asn Thr Asp Asp Leu Glu Asn Ser Ser Val Thr Ser Lys Gln
          50           55           60

Thr Pro Gln Leu Val Val Ser Glu Asp Pro Met Met Met Ser Ala Val
65           70           75           80

Pro Ser Ala Thr Ser Leu Asn Lys Ala Phe Ser Ile Asn Lys Glu Ile
          85           90           95

Gln Pro Gly Gln Ala Gly Leu Met Gln Thr Glu Arg Pro Gly Val Ser
          100          105          110

Thr Pro Thr Glu Ser Gly Val Pro Ser Ala Glu Glu Val Phe Gly Ser
          115          120          125

Ser Gln Pro Glu Arg Ile Ser Pro Glu Ser Gly Leu Ala Lys Ala Met
          130          135          140

Leu Thr Ile Ala Ile Thr Ala Thr Pro Ser Leu Thr Val Asp Glu Lys
145          150          155          160

Glu Glu Leu Leu Thr Ser Thr Asn Phe Gln Pro Ile Val Glu Glu Ile
          165          170          175

Thr Glu Thr Thr Lys Gly Phe Leu Lys Tyr Met Asp Asn Gln Ser Phe
          180          185          190

Ala Thr Glu Ser Gln Glu Gly Val Gly Leu Gly His Ser Pro Ser Ser
          195          200          205

```

Tyr Val Asn Thr Lys Glu Met Leu Thr Thr Asn Pro Lys Thr Glu Lys
 210 215 220
 Phe Glu Ala Asp Thr Asp His Arg Thr Thr Ser Phe Pro Gly Ala Glu
 225 230 235 240
 Ser Thr Ala Gly Ser Glu Pro Gly Ser Leu Thr Pro Asp Lys Glu Lys
 245 250 255
 Pro Ser Gln Met Thr Ala Asp Asn Thr Gln Ala Ala Ala Thr Lys Gln
 260 265 270
 Pro Leu Glu Thr Ser Glu Tyr Thr Leu Ser Val Glu Pro Glu Thr Asp
 275 280 285
 Ser Leu Leu Gly Ala Pro Glu Val Thr Val Ser Val Ser Thr Ala Val
 290 295 300
 Pro Ala Ala Ser Ala Leu Ser Asp Glu Trp Asp Asp Thr Lys Leu Glu
 305 310 315 320
 Ser Val Ser Arg Ile Arg Thr Pro Lys Leu Gly Asp Asn Glu Glu Thr
 325 330 335
 Gln Val Arg Thr Glu Met Ser Gln Thr Ala Gln Val Ser His Glu Gly
 340 345 350
 Met Glu Gly Gly Gln Pro Trp Thr Glu Ala Ala Gln Val Ala Leu Gly
 355 360 365
 Leu Pro Glu Gly Glu Thr His Thr Gly Thr Ala Leu Leu Ile Ala His
 370 375 380
 Gly Asn Glu Arg Ser Pro Ala Phe Thr Asp Gln Ser Ser Phe Thr Pro
 385 390 395 400
 Thr Ser Leu Met Glu Asp Met Lys Val Ser Ile Val Asn Leu Leu Gln
 405 410 415
 Ser Thr Gly Asp Phe Thr Glu Ser Thr Lys Glu Asn Asp Ala Leu Phe
 420 425 430
 Phe Leu Glu Thr Thr Val Ser Val Ser Val Tyr Glu Ser Glu Ala Asp
 435 440 445
 Gln Leu Leu Gly Asn Thr Met Lys Asp Ile Ile Thr Gln Glu Met Thr
 450 455 460
 Thr Ala Val Gln Glu Pro Asp Ala Thr Leu Ser Met Val Thr Gln Glu
 465 470 475 480
 Gln Val Ala Thr Leu Glu Leu Ile Arg Asp Ser Gly Lys Thr Glu Glu
 485 490 495
 Glu Lys Glu Asp Pro Ser Pro Val Ser Asp Val Pro Gly Val Thr Gln

500	505	510
Leu Ser Arg Arg Trp Glu Pro Leu Ala Thr Thr Ile Ser Thr Thr Val 515	520	525
Val Pro Leu Ser Phe Glu Val Thr Pro Thr Val Glu Glu Gln Met Asp 530	535	540
Thr Val Thr Gly Pro Asn Glu Glu Phe Thr Pro Val Leu Gly Ser Pro 545	550	555
Val Thr Pro Pro Gly Ile Met Val Gly Glu Pro Ser Ile Ser Pro Ala 565	570	575
Leu Pro Ala Leu Glu Ala Ser Ser Glu Arg Arg Thr Val Val Pro Ser 580	585	590
Ile Thr Arg Val Asn Thr Ala Ala Ser Tyr Gly Leu Asp Gln Leu Glu 595	600	605
Ser Glu Glu Thr Gly Phe His His Val Ala Gln Ala Arg Leu Lys Leu 610	615	620
Leu Gly Leu Arg Ser Leu Pro Ala Ser Ala Ser Gln Ser Val Gly Ile 625	630	635
Thr Ser Val Asn Ser Cys Thr Gln Pro Arg Lys Tyr Leu Asn Ser Cys 645	650	655
Leu Lys Trp Lys Leu Asn Pro Lys His Phe Ala Thr Gly Val 660	665	670

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1302 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CAGGATGGGA TCCTCATAGA GGGGCAGATA ATGTTTCTAC TAAATCTTCG GACAGTGATG	60
TAAGTGATAT ATCTGCGGTT TCAAGGACTA GTAGTGCTTC TCGTTTCAGC AGCACAAGCT	120
ACATGTCTGT CCAATCAGAA CGCCAGGAG GAAACAAGAA AATCAGTGTC TTTACATCCA	180
AAATGCAAAG CAGACAAATG GGCATATCAG GGAAGAACAT GACAAAAAGC ACCAGCATCA	240

```

GTGGAGACAT GTGCTCACTG GAGAAGAATG ATGGCAGCCA GTCTGACACT GCAGTGGGCA      300
CCTTGGGCAC CAGTGGCAAA AAGCGGCGCT CTAGCCTTGG TGCCAAAATG GTAGCTATCG      360
TTGGTCTGTC ACGGAAAAGT CGCAGTGCTT CTCAGCTCAG CCAAACGGAA GCAGGAGGTW      420
AAAAACTAAG GAGCACTGTC CAAAGAAGTA CAGAAACAGG CCTGGCCGTG GAAATGAGGA      480
ACTGGATGAC TCGACAGGCA AGCCGAGAGT CTACAGATGG TAGCATGAAC AGCTACAGCT      540
CAGAAGGAAA TCTGATTTTC CCTGGTGTTT GCTTGGCCTC TGATAGCCAG TTCAGTGATT      600
TCCTGGATGG CCTTGGCCCT GCTCAGCTAG TGGGACGCCA GACTCTGGCA ACACCTGCAA      660
TGGGTGACAT TCAGGTAGGA ATGATGGACA AAAAGGGACA GCTGGAGGTA GAAATCATCC      720
GGGCCCCGTG CCTTGTGTGA AAACCAGGTT CCAAGACACT GCCAGCACCG TATGTAAAAG      780
TGTATCTATT AGATAACGGA GTCTGCATAG CAAAAAAGGG ACAGCTGGAG GTAGAAATCA      840
TCCGGGCCCCG TGGCCTTGTT GTAAAACCAG GTTCCAAGAC ACTGCCAGCA CCGTATGTAA      900
AAGTGATCTT ATTAGATAAC GGAGTCTGSA TAGCCAAAAA GAAAACAAAA GTGGCAAGAA      960
AAACGCTGGA ACCCCTTTAC CAGCAGYTAT TATCTTTCGA AGAGAGTCCA CAAGGAAAAG     1020
TTTTACAGAT CATCGTCTGG GGAGATTATG GCCGCATGGA TCACAAATCT TTTATGGGAG     1080
TGGCCCAGAT ACTTTTAGAT GAACTAGAGC TATCCAATAT GGTGATCGGA TGGTTCAAAC     1140
TTTTCCACC TTCTCCCTA GTAGATCCAA CCTTGGCCCC TCTGACAAGA AGAGCTTCCC     1200
AATCATCTCT GGAAAGTTCA ACTGGACCTT CTACTYTCG TTCATAGCAG CTGTAAAAAA     1260
ATTGTTGTCA CAGCAACCAG CGTTACAAAA AAAAAAAAAA AA                          1302

```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met Ser Val Gln Ser Glu Arg Pro Gly Gly Asn Lys Lys Ile Ser Val
1           5           10           15
Phe Thr Ser Lys Met Gln Ser Arg Gln Met Gly Ile Ser Gly Lys Asn

```

73

Ala Gln Ile Leu Leu Asp Glu Leu Glu Leu Ser Asn Met Val Ile Gly
 325 330 335

Trp Phe Lys Leu Phe Pro Pro Ser Ser Leu Val Asp Pro Thr Leu Ala
 340 345 350

Pro Leu Thr Arg Arg Ala Ser Gln Ser Ser Leu Glu Ser Ser Thr Gly
 355 360 365

Pro Ser Tyr Xaa Arg Ser
 370

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1627 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

GGCCTACTGA ATAACAATAG TTATTTTGA AAGATCATTG TTTCCCTAAT TCTGTGGCTA      60
AGAGGGGCCC ATTTAGGAAG AGGTCAGGAG GGTGGTTGAG GCGGGGAGGG GGA CTCTTAA      120
GGAAAAGACA ACTCAAACGG GAGGATGAAA AAGCATTGTA TTTGATTGTT TTCTTTCTTG      180
TAGGATGATG CCTCCCAGGA GAATAACATG CAAGCTGAGG TGACAGGTAA ACTAAAACAG      240
CACTTTAATA ATGCGCTCCC CAAACAGAAC AGCACTACCA CCGGGGTCTC CGTCATCATG      300
GTGAGTGGAA GCGGCAGGT CTGGCCTGAC TCGGAGCCG GCCTTGAAGT TTTTGAATTA      360
GGTAGCCGGG AGCTGCCCTC ACATGGAAGT TGGTGCCTTC CGTAGTCCTA TTTCATATGA      420
AGATTGGCTT GGCATGTGGA GGGCACTCAT TCGGCAACTC CCAGGCTTTG GGCACTGTGT      480
GGAGGGGCTT GTGTAGGGAC CAGCAGGCCT GGTGTGAGGG GTCCAGGCGT CAAGGAGCTC      540
CTGGCTGGGC CCTCTGGGCA GCTGCTTCCA CTC TTGTCTC TGCCTTCTCA TCTAGAGAGA      600
CTCCAAGCC CTGGAGGGGT GTGTTGTGTT AGGAATTAAC TCCCTGCCTA CCCCAGGCC      660
TCAGAAATAG ATTATTAGAG ATGTGAATTA TTC TTGAGA CTTGGGATAA GAAACAGCCA      720
AAGCTAAACA TATTTCACTT TAAAAAATC AGTGT TTAT AAAACACAGT TTGGGGCTTT      780
TAAAGGTACA TAATCAAGGA AAAAAATATA TATTCATTTT TCAGGGTTGG TAACATTTTA      840

```

```

TGAGATGTCA GTGACAACGA TGGCCTTATT TTTTTCAGCC TTTTCTTCTT CCAAAATGTT      900
TCTTAAGGCA ACTCTCCTAA ATACATAAAC ACAACAAATT AAAATGAAAA GTGACATGAG      960
AGTAAATGAA TCAAAAGGAA AAAACATTGA ACCAGAGGTG AGGGCAGCAC ACCCGCAGCA      1020
GCTGTCCAGG CCTGAGCCAA TGCAACCCTG GCGGGGAAGG CCAGCTCACC GTGAGCAGGT      1080
AGAAGCCAGC CAGCCACCCA GGCAGGGACC TTGGTTCTCC CCACACACTC CCAGGAGCAG      1140
GGAACAGGGG TGGAGTGGCC TTTCCAGAG CTGGAGTTGG CTGCAGCAGC TTTCGAATCA      1200
GACCTGCCAA GGTGATGGGC GTCTGAGTTT CACATCTGGG CCCCCGTGA CCCCACTGAG      1260
TCCTGACAGC TAAGGATGGG CCACCTCCAC AGCTCCGTCA CTCGTACTTG GGACAGGCCT      1320
CTCATCCTCT GGAAGGTCC TCCTTGTTTC CTACCCAACT AGAAGGGAAA CAGTGGCATA      1380
TTCTCATGGT ACATGGTTGT CTGAAAGCCT TACCTAGGAA GACGCAGGGT CTAGATAGAA      1440
GCTATAAGGA AGCCACACAC ATAACCCACA TCCCCACACC CCCAACATCC CCCCACTCC      1500
CCACACCCCC CACACCCCCC ACATCCCCAC CATAATTACC CCCACCTCCA AATATCTCAT      1560
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA      1620
AAAAAAA                                         1627

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 65 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Met Gln Ala Glu Val Thr Gly Lys Leu Lys Gln His Phe Asn Asn Ala
1           5           10           15
Leu Pro Lys Gln Asn Ser Thr Thr Thr Gly Val Ser Val Ile Met Val
20           25           30
Ser Gly Arg Arg Gln Val Trp Pro Asp Cys Gly Ala Gly Leu Glu Val
35           40           45
Phe Glu Leu Gly Ser Arg Glu Leu Pro Ser His Gly Ser Trp Cys Leu
50           55           60

```

Pro
65

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2468 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

AAATTTTCAT CTCTAAATCA AAAAGTGAAG CAGACCAAAT CCAATGTAAA TATTGGCAAC      60
CTCCGAAAGC TAGGAAACTT TACCAAACCT GAAATGAAAG TTAACTTTCT AAAACCAAAC      120
TTAAAAGTAA ATCTTTGGAA ATCAGATAGT AGTCTTGAAA CTATGGAAAA CACAGGAGTG      180
ATGGATAAGG TTCAGGCAGA GTCTGATGGG GACATGTCTT CAGATAATGA CTCATACCAC      240
TCTGATGAAT TCCTTACAAA TTCTAAGTCT GATGAAGACA GGCAGCTAGC TAACTCATT      300
GAGAGTGTAG GGCCAATAGA TTACGTTCTT CCTAGTTGTG GTATTATTGC CTCAGCGCCT      360
CGATTGGGCA GTCGGTCCCA GTCTCTTAGC AGCACAGATA GTAGCGTTCA TGCTCCTTCA      420
GAGATTACTG TTGCTCATGG GAGTGGGCTT GGAAAAGGCC AGGAGTCTCC TTTGAAGAAA      480
AGTCCTTCTG CTGGCGACGT ACACATATTG ACTGGCTTTG CCAAGCCTAT GGATATTTAC      540
TGCCACAGAT TTGTGCAAGA TGCACAGAAC AAAGTGACCC ACCTATCAGA GACCAGATCT      600
GTGTCTCAGC AGGCTAGTCA GGAAAGAAAT CAAATGACCA ATCAAGTTTC AAATGAAACC      660
CAATCAGAAAT CAACAGAACA GACACCTTCT CGGCCATCGC AATTAGATGT CTCTCTTTCT      720
GCAACAGGCC CACAGTTTTT GTCAGTTGAG CCAGCGCATT CAGTTGCATC TCAAAAAACC      780
CCCACCTCCG CTTCCAGCAT GCTTGAACCT GAGACAGGGC TTCATGTAAC TCCTTCTCCT      840
TCAGAGAGCA GTAGCAGCAG AGCAGTCTCT CCCTTTGCCA AGATTGGAAG TTCCATGGTC      900
CAGGTTGCTA GTATTACCCA AGCTGGATTG ACCCATGGGA TAACTTTGC AGTGTCAAAA      960
G TTCAGAAGA GTCCTCCAGA ACCTGAAATC ATTAATCAAG TCCAGCAAAA TGAAC TTAAA      1020
AAGATGTTTA TACAATGCCA GACACGGATA ATTCAGATTT AGCTTTTAGC CATAAGAATC      1080
CTTCCATGGC TTTTATTTAA AAATATGAAA TTTTCACCTC TTGGGGTATT TTAATGTAC      1140

```

TGTCTGAACC CAGGGATCAC AAATTCGTGT CATTGGAAAG GGTTTTAAAC GGAGTCGGAA	1200
CCTGAGTAGA TTTCCAAATTTTACAGCCAG GACTACAGAA GTGCATCATT CTAGAATGTG	1260
TAGACCTGAG TAGCTTATAC ACTACAGAGC ACTTTGCTTA TTTGAAAGTA ATTCAGCAAC	1320
AGGTCACCTTT GGGATATAAC CTGAACCTTT TTTTGGAGTG GGGTGGGTAG ACTACAGTAG	1380
ACACAAGGGC TGGACATGCA GATGCTTAGG GGATTAGCGT TTTTCATAAT TTGTTCTGTT	1440
TGTCAGTTCA TTCCTGTGTG TTCTTACCTC TACAAAGTAA ATTACACATT TAGTTTTTAG	1500
TGACTTTAAC ATGTTACTGA AGCATTTGAA TATAAAGCTA TTTTAGTTTT GATGGCTTAA	1560
CTGTTCCCTG AGGAGTTGAG GGTATTGAC ACTAAAAAA TGAATTCTCA TTTGATCCTA	1620
ATTTTCCCCG TATTCTACTT GAACACATTA AAAATACTCT GCTGCCTATA CAATGTAAAC	1680
CTAGGAGCAT TAAGACTTGT CACACAGTAA ACCTGATACA TCAGAGGTGA ATACCAGCAC	1740
CTATTAGGTT TCATTTTGCT GTTTTCAGGA ATGTAAGAAC ACCCATATTG GCTACTGGAA	1800
ATTCTAGCAG TCAGTCAGGT TTTAAATTTAT TCCAGGAGGG GCATCCTCGA CATCTTATGT	1860
AGATGATCCA CAACTTCAAA ATTTAGTCTG GGCTTAGTGC AGTGGCTCAC ACCTATAATC	1920
CCAACACTTT GGGAGGCCAG GAGTTTGTAG ACCAGCCTGG GAAACATCTG CCTCTACAAA	1980
AAAATACAAA AATTAGCTGG GCATAGTGGT GCATGCCTGT GTTCTAGCT ACGCAGGAGG	2040
ATTGCTTGAG CCCATGAGAT TGAGGCTGCA GTGAGCTGTG ATCGTGCCAC TGACCTCCAG	2100
CCTGGGGGAC AGAGCAAGAC CGTGTCTCAA AAACAATTTA GTCTGAAACA CAATTGTGCT	2160
GAATCTGTCT GACTATAACT CTGACCACAC AGAACCAGGG CTGCCCCGTG AATCCCCACA	2220
GTAAGAAAGT TGTATGGCAT ATTCCAACAA GTATTGGTTC GTCTGGTGTC TTTAGAGCTT	2280
TACTCTGTTG AAGTACTGA TTCTCAACTG AACATTATGT CGTTACTTTG ATAAGCATTC	2340
CACTTTTGTT ATTTATTAGT GCTATCTTTT TTTTPTTACG TGTTAAATCT TGTGATTATT	2400
AAAATAAAGT ACCATTGTAA TTTAAAGTGA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	2460
AAAAAAAA	2468

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 322 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Lys	Val	Asn	Phe	Leu	Lys	Pro	Asn	Leu	Lys	Val	Asn	Leu	Trp	Lys	1	5	10	15
Ser	Asp	Ser	Ser	Leu	Glu	Thr	Met	Glu	Asn	Thr	Gly	Val	Met	Asp	Lys	20	25	30	
Val	Gln	Ala	Glu	Ser	Asp	Gly	Asp	Met	Ser	Ser	Asp	Asn	Asp	Ser	Tyr	35	40	45	
His	Ser	Asp	Glu	Phe	Leu	Thr	Asn	Ser	Lys	Ser	Asp	Glu	Asp	Arg	Gln	50	55	60	
Leu	Ala	Asn	Ser	Leu	Glu	Ser	Val	Gly	Pro	Ile	Asp	Tyr	Val	Leu	Pro	65	70	75	80
Ser	Cys	Gly	Ile	Ile	Ala	Ser	Ala	Pro	Arg	Leu	Gly	Ser	Arg	Ser	Gln	85	90	95	
Ser	Leu	Ser	Ser	Thr	Asp	Ser	Ser	Val	His	Ala	Pro	Ser	Glu	Ile	Thr	100	105	110	
Val	Ala	His	Gly	Ser	Gly	Leu	Gly	Lys	Gly	Gln	Glu	Ser	Pro	Leu	Lys	115	120	125	
Lys	Ser	Pro	Ser	Ala	Gly	Asp	Val	His	Ile	Leu	Thr	Gly	Phe	Ala	Lys	130	135	140	
Pro	Met	Asp	Ile	Tyr	Cys	His	Arg	Phe	Val	Gln	Asp	Ala	Gln	Asn	Lys	145	150	155	160
Val	Thr	His	Leu	Ser	Glu	Thr	Arg	Ser	Val	Ser	Gln	Gln	Ala	Ser	Gln	165	170	175	
Glu	Arg	Asn	Gln	Met	Thr	Asn	Gln	Val	Ser	Asn	Glu	Thr	Gln	Ser	Glu	180	185	190	
Ser	Thr	Glu	Gln	Thr	Pro	Ser	Arg	Pro	Ser	Gln	Leu	Asp	Val	Ser	Leu	195	200	205	
Ser	Ala	Thr	Gly	Pro	Gln	Phe	Leu	Ser	Val	Glu	Pro	Ala	His	Ser	Val	210	215	220	
Ala	Ser	Gln	Lys	Thr	Pro	Thr	Ser	Ala	Ser	Ser	Met	Leu	Glu	Leu	Glu	225	230	235	240
Thr	Gly	Leu	His	Val	Thr	Pro	Ser	Pro	Ser	Glu	Ser	Ser	Ser	Ser	Arg	245	250	255	
Ala	Val	Ser	Pro	Phe	Ala	Lys	Ile	Arg	Ser	Ser	Met	Val	Gln	Val	Ala	260	265	270	

Ser Ile Thr Gln Ala Gly Leu Thr His Gly Ile Asn Phe Ala Val Ser
 275 280 285

Lys Val Gln Lys Ser Pro Pro Glu Pro Glu Ile Ile Asn Gln Val Gln
 290 295 300

Gln Asn Glu Leu Lys Lys Met Phe Ile Gln Cys Gln Thr Arg Ile Ile
 305 310 315 320

Gln Ile

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1044 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGGCCGCGG TGGTCCGGGC CGCCGTGTCT GCTTCCAGAG CCAATCTGC CGAGGCTGGA 60

ATTGCCGGGG AGGCCCAAAG CAAGAAGCCA GTGTCCAGGC CGGCCACCGC TGCCGCTGCC 120

GCTGCCGGCT CCAGGGAGCC CCGTGTCAAG CAAGGTCCAA AAATTTATAG TTTTAATTCT 180

ACAAATGATT CTAGTGGTCC TGCAAATCTG GATAAATCTA TTTTGAAAGT GGTAATTAAT 240

AACAACTAG AGCAAAGAAT TATTGGAGTG ATCAATGAGC ATAAAAAGCA AAATAATGAC 300

AAAGGAATGA TTTCTGGAAG ACTTACTGCC AAAAAATTGC AGGATTTATA CATGGCTTTA 360

CAAGCATTTT CATTTAAGAC AAAGGACATT GAAGATGCCA TGACCAATAC ACTCTTATAT 420

GGAGGTGACC TTCATTCTGC CTTGGATTGG CTCTGTTTAA ACCTTTCAGA TGATGCACTT 480

CCTGAAGGAT TCAGTYAGGA ATTWGAAGAG CAGAAACCTA AAAGTAGGCA TAAATTTTCA 540

TCTCCTCAAA TATAAGTTAC TATTTACCTT CCATTGCAAC CTAAAACAAA AACATATGAA 600

GAGGACCCTA AGAGTAAGCC AAAAAAGGAA GAAAAAATA TGAAGTAAA TATGAAAGAG 660

TGGATTTTAC GATATGCTGA ACAACAAAAT GAAGAAGAAA AGAATGAGAA TTCTAAAAGT 720

TTAGAAGAGG AGGAAAAATT TGACCCTAAT GAAAGGTACT TACATCTTGC AGCAAACTG 780

CTGGATGCAA AAGAACAAGC AGCTACCTTT AACTAGAAA AAAACAAGCA AGGCCAAAAA 840

GAGGCTCAGG AAAAAATAAG GAAATTTCAA AGAGAAATGG AAACCTTTAGA AGACCATCCA 900
 GTATTTAACC CAGCCATGAA GATTTACAT CAACAAAATG AAAGGAAAAA GCCTCCTGTA 960
 GCCACAGAAG GAGAAAGTGC ATTGAATTTT AATTTATTTG AAAAATCTGC AGCTGCTACT 1020
 GAAGAAGAGA AAAAAAAAAA AAAA 1044

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Glu Val Asn Met Lys Glu Trp Ile Leu Arg Tyr Ala Glu Gln Gln
 1 5 10 15
 Asn Glu Glu Glu Lys Asn Glu Asn Ser Lys Ser Leu Glu Glu Glu
 20 25 30
 Lys Phe Asp Pro Asn Glu Arg Tyr Leu His Leu Ala Ala Lys Leu Leu
 35 40 45
 Asp Ala Lys Glu Gln Ala Ala Thr Phe Lys Leu Glu Lys Asn Lys Gln
 50 55 60
 Gly Gln Lys Glu Ala Gln Glu Lys Ile Arg Lys Phe Gln Arg Glu Met
 65 70 75 80
 Glu Thr Leu Glu Asp His Pro Val Phe Asn Pro Ala Met Lys Ile Ser
 85 90 95
 His Gln Gln Asn Glu Arg Lys Lys Pro Pro Val Ala Thr Glu Gly Glu
 100 105 110
 Ser Ala Leu Asn Phe Asn Leu Phe Glu Lys Ser Ala Ala Ala Thr Glu
 115 120 125
 Glu Glu
 130

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 764 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

AGGTC'TAGAA TTTGTTATTT TAAGTTGAAA TTTGAAGATT GGCTGAGTTG GTTTCATGAT      60
ATTGTTGCCT ACTGATATCT TACACAGAGT ACGAGTACAT CTGGTTGAAA GAGAAAAGGC      120
AGAGAGAAAA GCAGACTTTT TGGAGTTGTT GTTCTCAGGC ACTAACAGAT TCATCCCTTA      180
TATGCATATG AATGGACAAT AAGATTTCTT GGATATTTTC ATGGCAAAGG TGAGGAGAAA      240
TTTTTATATGC TTTTGAAGA AACATGATTT TTATTACCTT GGGTTATTC AGGAGATAAT      300
TGATAAATGT TGAGTAGTTT GTTGGTCTTC TTCTTAAAGG AGGATCAAGA GAAGCAGTAC      360
AGCCTGGGTG ACAGAGCGAG ACTCTGTCTC AAATAAAAAA AAAGAAGGAA ACATACTGTA      420
GCCAGACGCC ACCTATAGTC CCAAGTAGCT GGGATTACAG GACAAAACAA TTGAAGAGCT      480
CAGGCAATCT TTAGCAAATG TTGAAAGGAT GAAAGAGAAG GCAAATGTTG AAACGATGAA      540
AGAGAAGGCA GTTGTGAAAA CAGAAACTT GAAACTACA TTAGACTCTG CAGAGCAAAA      600
GGCAAGATCA GACAAAGAGA AGACCCAGCA GATGTTAGAT GCTGTCACCT CTGAGCCCCC      660
AACAGCAAAG AGCGCACCTG AAGAAGTATC AGGACAAGAA CAAGAGGTTT TTTCAAAATA      720
GTAAATTTAA AATTAATTTA GTTGAAGTGA TAAAAA AAAA                          764

```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 70 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met Lys Glu Lys Ala Asn Val Glu Thr Met Lys Glu Lys Ala Val Val
1           5           10          15

Lys Thr Glu Asn Leu Lys Thr Thr Leu Asp Ser Ala Glu Gln Lys Ala

```

	20		25		30
Arg Ser Asp Lys Glu Lys Thr Gln Gln Met Leu Asp Ala Val Thr Ser					
35		40		45	
Glu Pro Pro Thr Ala Lys Ser Ala Pro Glu Glu Val Ser Gly Gln Glu					
50		55		60	
Gln Glu Val Phe Ser Lys					
65		70			

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 666 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

GCGAGTCTGT CGATCCCAGG CCAGAGACAA GGCAGACTAA GGTTCATTG TAAAGAAGCT      60
CCTTCCAGCA CCTCCTCTCT TCTCCTTTTG CCCAACTCA CCCAGTGAGT GTGAGCATTT      120
AAGAAGCATC CTCTGCCAAG ACCAAAAGGA AAGAAGAAAA AGGGCCAAAA GCCAAAATGA      180
AACTGATGGT ACTTGTTTTT ACCATTGGGC TAACTTTGCT GCTAGGAGTT CAAGCCATGC      240
CTGCAAATCG CCTCTCTTGC TACAGAAAGA TACTAAAAGA TCACAACTGT CACAACCTTC      300
CGGAAGGAGT AGCTGACCTG ACACAGATTG ATGTCAATGT CCAGGATCAT TTCTGGGATG      360
GGAAGGGATG TGAGATGATC TGTTACTGCA ACTTCAGCGA ATTGCTCTGC TGCCCCAAAG      420
ACGTTTTCTT TGGACCAAAG ATCTCTTTTC TGATTCCTTG CAACAATCAA TGAGAATCTT      480
CATGTATTCT GGAGAACACC ATTCCTGATT TCCCACAAAC TGCACTACAT CAGTATAACT      540
GCATTTCTAG TTTCTATATA GTGCAATAGA GCATAGATTC TATAAATTCT TACTTGTCTA      600
AGACAAGTAA ATCTGTGTTA AACAAGTAGT AATAAAAGTT AATTCAATCT AAAAAAAAAA      660
AAAAAA

```

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 98 amino acids
 - (B) TYPE: amino acid

- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

Met Lys Leu Met Val Leu Val Phe Thr Ile Gly Leu Thr Leu Leu Leu
1           5           10           15

Gly Val Gln Ala Met Pro Ala Asn Arg Leu Ser Cys Tyr Arg Lys Ile
          20           25           30

Leu Lys Asp His Asn Cys His Asn Leu Pro Glu Gly Val Ala Asp Leu
          35           40           45

Thr Gln Ile Asp Val Asn Val Gln Asp His Phe Trp Asp Gly Lys Gly
          50           55           60

Cys Glu Met Ile Cys Tyr Cys Asn Phe Ser Glu Leu Leu Cys Cys Pro
65           70           75           80

Lys Asp Val Phe Phe Gly Pro Lys Ile Ser Phe Val Ile Pro Cys Asn
          85           90           95

Asn Gln

```

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CNAAGCATATA AGTCTCAGGG GCAGGTT

28

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNATTGCTGTG TAACCAATGC AAAACTGA

29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CNCAGAAGAT CCAGGACTAT AAACAGGAT

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CNGACAAATGT GCAATACAAT CGGTCCTC

29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GNCCATTTGTC TGCTTTGCAT TTTGGATG

29

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CNGCTTGCATG TTATTCTCCT GGGAGGCA

29

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CNATCACTCCT GTGTTTTCCA TAGTTTCA

29

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GNTCTTTTGCA TCCAGCAGTT TTGCTGCA

29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CNACATTTGCC TTCTCTTCA TCCTTCA

29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TNAACTCCTAG CAGCAAAGTT AGCCCAAT

29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 151 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```

Met Ser Gln Asn Met Gly Cys Tyr Phe Leu Phe Leu Ser Trp Phe Pro
1          5          10          15

Tyr Ser His Pro Ser Phe Ser Thr Gln Val Ser Gly Gln Glu Asp Phe
          20          25          30

Ser His Gln Leu Tyr Gln Arg Lys Leu Gln Ala Pro Leu Trp Pro Ser
          35          40          45

Ser Leu Gly Ile Thr Asp Cys Cys Gln Tyr Val Thr Ser Cys His Pro
          50          55          60

Lys Arg Ser Glu Arg Arg Lys Tyr Gly Arg Asp Phe Leu Leu Arg Phe
65          70          75          80

Arg Phe Cys Ser Ile Ala Cys Gln Arg Pro Val Gly Leu Val Leu Met
          85          90          95

Glu Gly Val Thr Asp Thr Lys Pro Glu Arg Pro Ala Gly Trp Ala Glu
          100          105          110

Ser Val Leu Glu Glu Asp Ala Ser Glu Leu Glu Pro Ala Phe Ser Arg
          115          120          125

Thr Val Gly Thr Ile Gln His Cys Leu His Leu Thr Ser Val Tyr Thr
          130          135          140

Pro Ser Thr Pro Ser Arg Ala
145          150

```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 71 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met	Lys	Ile	Gly	Leu	Ala	Cys	Gly	Gly	His	Ser	Phe	Gly	Asn	Ser	Gln
1				5					10					15	
Ala	Leu	Gly	Thr	Val	Trp	Arg	Gly	Leu	Cys	Arg	Asp	Gln	Gln	Ala	Trp
			20					25					30		

Cys Glu Gly Ser Arg Arg Gln Gly Ala Pro Gly Trp Ala Leu Trp Ala
35 40 45

Ala Ala Ser Thr Leu Val Ser Ala Phe Ser Ser Arg Glu Thr Pro Lys
50 55 60

Pro Trp Arg Gly Val Leu Cys
65 70

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:
 - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
 - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 101 to nucleotide 428;
 - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 325 to nucleotide 461;
 - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AA365_1 deposited under accession number ATCC 98296;
 - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296;
 - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AA365_1 deposited under accession number ATCC 98296;
 - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296;
 - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
 - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
 - (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
 - (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
 - (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).
2. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.
3. A host cell transformed with a composition of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein encoded by a composition of claim 2, which process comprises:
 - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
 - (b) purifying said protein from the culture.
6. A protein produced according to the process of claim 5.
7. The protein of claim 6 comprising a mature protein.
8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:
 - (a) the amino acid sequence of SEQ ID NO:2;
 - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 76 to amino acid 109;
 - (c) fragments of the amino acid sequence of SEQ ID NO:2; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone AA365_1 deposited under accession number ATCC 98296;the protein being substantially free from other mammalian proteins.
9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 76 to amino acid 109.
11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.
12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.
14. A composition comprising an isolated polynucleotide selected from the group consisting of:
 - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
 - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 245 to nucleotide 421;
 - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 413 to nucleotide 421;
 - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BL67_2 deposited under accession number ATCC 98296;
 - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BL67_2 deposited under accession number ATCC 98296;
 - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BL67_2 deposited under accession number ATCC 98296;
 - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BL67_2 deposited under accession number ATCC 98296;
 - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
 - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;
 - (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
 - (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
 - (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).
15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:
 - (a) the amino acid sequence of SEQ ID NO:4;

- (b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 20;
 - (c) fragments of the amino acid sequence of SEQ ID NO:4; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone BL67_2 deposited under accession number ATCC 98296;
- the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3 or SEQ ID NO:5.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 137 to nucleotide 2146;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 218 to nucleotide 2146;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 189 to nucleotide 397;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CG99_2 deposited under accession number ATCC 98296;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CG99_2 deposited under accession number ATCC 98296;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CG99_2 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CG99_2 deposited under accession number ATCC 98296;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 50 to amino acid 87;

(c) fragments of the amino acid sequence of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone CG99_2 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 123 to nucleotide 1247;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 1 to nucleotide 419;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO618_1 deposited under accession number ATCC 98296;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO618_1 deposited under accession number ATCC 98296;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 99;

(c) fragments of the amino acid sequence of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO618_1 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 208 to nucleotide 402;

- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 184 to nucleotide 300;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO629_2 deposited under accession number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO629_2 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 31;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CO629_2 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

26. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 94 to nucleotide 1059;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 387;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT645_1 deposited under accession number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT645_1 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:16;
- (b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 98;

- (c) fragments of the amino acid sequence of SEQ ID NO:16; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone CT645_1 deposited under accession number ATCC 98296;
- the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 640 to nucleotide 1029;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 611 to nucleotide 814;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CW383_1 deposited under accession number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CW383_1 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
 - (b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 63;
 - (c) fragments of the amino acid sequence of SEQ ID NO:18; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone CW383_1 deposited under accession number ATCC 98296;
- the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 509 to nucleotide 718;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN167_1 deposited under accession number ATCC 98296;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN167_1 deposited under accession number ATCC 98296;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN167_1 deposited under accession number ATCC 98296;
- (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN167_1 deposited under accession number ATCC 98296;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and
- (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:20;
- (b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 65;
- (c) fragments of the amino acid sequence of SEQ ID NO:20; and
- (d) the amino acid sequence encoded by the cDNA insert of clone DN167_1 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 177 to nucleotide 470;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 237 to nucleotide 470;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN711_2 deposited under accession number ATCC 98296;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN711_2 deposited under accession number ATCC 98296;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:22 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:22;

(b) the amino acid sequence of SEQ ID NO:22 from amino acid 1 to amino acid 94;

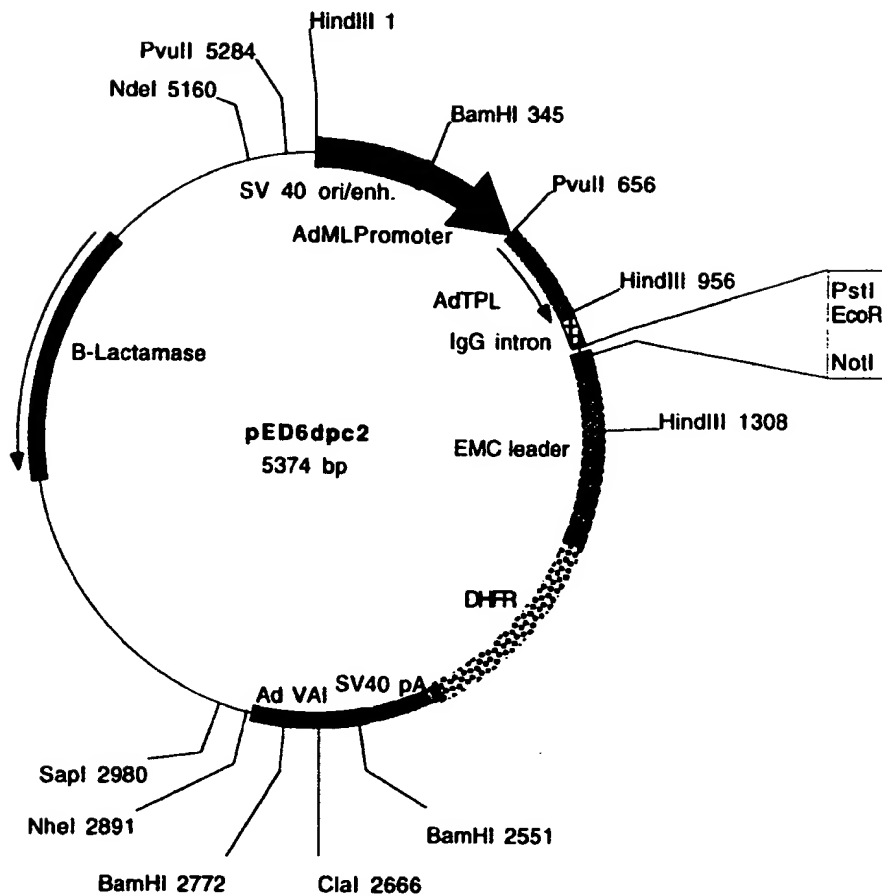
(c) fragments of the amino acid sequence of SEQ ID NO:22; and

(d) the amino acid sequence encoded by the cDNA insert of clone DN711_2 deposited under accession number ATCC 98296;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:21.

FIGURE 1A

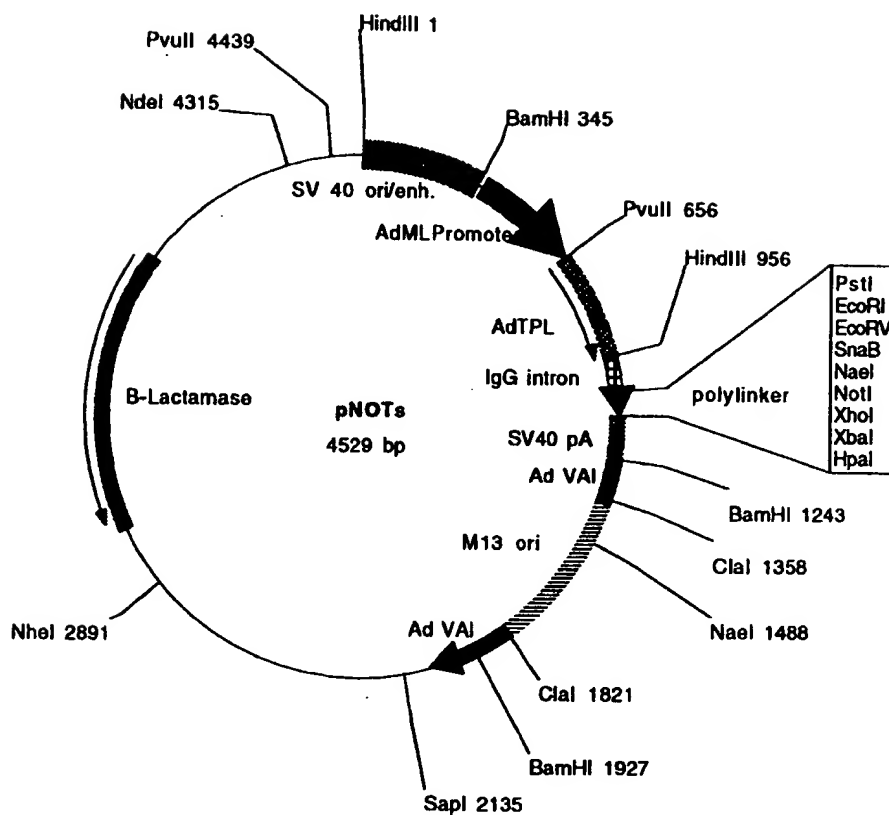


Plasmid name: pED6dpc2

Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs

Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI